REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-01-0188*

1. REPORT DATE *(DD-MM-YYYY)*
08–2002

2. REPORT TYPE
Technical

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE

NON-LINEAR EFFECTS IN ADAPTIVE LINEAR PREDICTION

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0601152N

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHORS

A. A. (Louis) Beex
Virginia Tech

J. R. Zeidler
SSC San Diego

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

SSC San Diego
San Diego, CA 92152-5001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
800 North Quincy Street
Arlington, VA 22217–5000

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

20090803051

13. SUPPLEMENTARY NOTES

14. ABSTRACT

When a conventional normalized least mean square (NLMS) adaptive filter is used to predict a process, especially when predicting several samples ahead, non-linear effects can be observed. These non-linear effects produce adaptive filter performance that exceeds that of the conventional Wiener filter, and engenders weight behavior that is of a time-varying nature. After showing the existence of such non-linear effects, we show their relation to the difference between the structure of the optimal predictor and the structure used to model the data to be predicted. The non-linear effects are stronger when the process to be predicted is more narrowband.

Published in *Proceedings of the 4th IASTED International Conference on Signal and Image Processing*, 21–26.

15. SUBJECT TERMS
Mission Area: Communications

non-linear effects	multi-channel Wiener filter	adaptive prediction
time-varying Wiener filter	multi-channel adaptive filter	normalized least mean square

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19B. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	6	J. R. Zeidler / (619) 553-1581

Non-Linear Effects in Adaptive Linear Prediction

A. A. (Louis) Beex & James R. Zeidler
DSPRL – ECE 0111
Virginia Tech
Blacksburg, VA 24061-0111, USA

Communications & Information Systems, 28505
SPAWAR Systems Center
San Diego, CA 92152, USA

ABSTRACT

When a conventional NLMS adaptive filter is used to predict a process, especially when predicting several samples ahead, non-linear effects can be observed. These non-linear effects produce adaptive filter performance that exceeds that of the conventional Wiener filter, and engenders weight behavior that is of a time-varying nature. After showing the existence of such non-linear effects, we show their relation to the difference between the structure of the optimal predictor and the structure used to model the data to be predicted. The nonlinear effects are stronger when the process to be predicted is more narrowband.

KEY WORDS: non-linear effects, NLMS, time-varying Wiener filter, multi-channel Wiener filter, multi-channel adaptive filter, adaptive prediction.

1. INTRODUCTION

Non-linear effects have been shown to exist in a number of adaptive filtering applications, such as adaptive noise canceling [1, 2], interference contaminated adaptive equalization [1, 3], and adaptive linear prediction of chirped processes [4]. Often, though not exclusively, these nonlinear effects are more prominent when bandwidths are narrow and when adaptive filter stepsizes are relatively large. The nonlinear effects are characterized by performance that exceeds that of the Wiener filter of the same structure, and by time-varying adaptive filter weight behavior. In adaptive linear prediction of chirped processes the performance depends on chirp rate and bandwidth [4].

In an adaptive linear prediction scenario with a wide-sense stationary AR(1) process, it was shown [5] that the non-linear effect is stronger the farther ahead one aims to predict. The latter means that the loss in prediction performance, associated with the increase in prediction distance, is less for the adaptive filter than for the Wiener filter of the corresponding structure. These results are especially important for applications such as the prediction of narrowband data in correlated wideband noise where the selection of a prediction distance that exceeds the correlation length of the additive noise can enhance the predictability of the narrowband component.

While nonlinear effects in adaptive filtering have been shown to exist with the least-mean-square (LMS) algorithm as well as its normalized form (NLMS), we will concentrate here on using the NLMS algorithm. We will begin by showing that nonlinear effects exist in the adaptive linear prediction (ALP) scenario. As was shown for the noise canceling, equalizer, and prediction contexts [1 - 4], fundamentally the nonlinear effects originate from the error signal feedback, which is used in the weight update of the NLMS algorithm.

Here we aim to reveal the mechanism by which the error feedback results in the observed nonlinear effects. The error signal carries instantaneous information about the discrepancy between the actual desired data and its NLMS modeled version. The latter is thus related to the structure that underlies the optimal estimator for the ALP scenario being investigated. We will see that the structure of the optimal estimator is different from the tapped delay line structure used in conventional adaptive filtering. Forcing the conventional tapped delay line model to identify the structure of the optimal estimator results in an equivalent tapped delay line structure with time-varying weights. NLMS tracking of the latter can produce the performance gain associated with the observed nonlinear effects.

In addition we illustrate that the performance of the adaptive linear predictor is bounded by a two channel Wiener filter that utilizes the conventional reference channel, containing samples of the far past of the input process, and a second or auxiliary channel containing samples of the most recent past of the input channel. A two-channel Wiener filter such as that considered here could be implemented in an approximate fashion by using the far past inputs to estimate the most recent past, and then using the latter as the second channel.

2. ALP SCENARIO

In the adaptive linear prediction (ALP) scenario of interest, the process to be predicted is a white noise n_n contaminated autoregressive process of order 1. We will limit ourselves here to a first order AR(1) process s_n, in view of the fact that the nonlinear effects reported to date involved AR(1) processes. The process to be predicted is therefore d_n.

$$d_n = s_n + n_n \tag{1}$$

A causal linear predictor, for predicting Δ steps ahead, would use a linear combination of samples of the desired signal, available at time *n*, to predict $d_{n+\Delta}$.

$$\begin{aligned} \hat{d}_{n+\Delta} &= l.c.\{d_m\}_{m=-\infty}^{n} \\ &= \sum_{k=0}^{\infty} h_k d_{n-k} \end{aligned} \tag{2}$$

In preparation for the adaptive filtering context we introduce a delay of Δ samples into (2), because adaptation will be done on the basis of the error at time *n*. For a wide sense stationary process, the resulting predictor would remain the same, so that we have the following.

$$\begin{aligned} \hat{d}_n &= l.c.\{d_m\}_{m=-\infty}^{n-\Delta} \\ &= \sum_{k=0}^{\infty} h_k d_{n-\Delta-k} \end{aligned} \tag{3}$$

The unit pulse response of an optimal linear predictor, as in (2) and (3), is of infinite length. In the adaptive linear predictor there is a limit to the number of unit pulse response samples that can be used, let's say *M*. The output of the adaptive linear predictor is therefore represented as follows.

$$\begin{aligned} y_n &= l.c.\{d_m\}_{m=n-\Delta-M+1}^{n-\Delta} \\ &= \sum_{k=0}^{M-1} w_{k,n}^{*} d_{n-\Delta-k} \\ &= \mathbf{w}_n^H \mathbf{r}_n \end{aligned} \tag{4}$$

We have indicated explicitly that the adaptive filter weights vary during adaptation. The samples of the process used in forming the adaptive filter prediction y_n, are contained in $\mathbf{r}_n$, the reference input vector.

$$\mathbf{r}_n = \begin{bmatrix} d_{n-\Delta} \\ d_{n-\Delta-1} \\ \vdots \\ d_{n-\Delta-(M-1)} \end{bmatrix} \tag{5}$$

Since all samples are delayed by Δ or more, we refer to the reference vector as containing the far past of the process.

As far as prediction goes, prediction from the nearest past results in better performance for this white noise input. For additive correlated noise it is often desirable to delete the near past, which has strong noise correlation, in favor of larger prediction distances, at which the noise components are uncorrelated with the current data. For purposes of comparison, and – as we will see – explanation, an auxiliary input vector $\mathbf{x}_n$ is defined. The latter contains the most recent past of the process to be predicted, as expressed in the following definition.

$$\mathbf{x}_n = \begin{bmatrix} d_{n-1} \\ d_{n-2} \\ \vdots \\ d_{n-(L-1)} \end{bmatrix} \tag{6}$$

The adaptive filtering scenario, using either the reference vector only (the conventional case) or the auxiliary vector in addition (the 2-channel case), can then be represented as in Fig. 1.

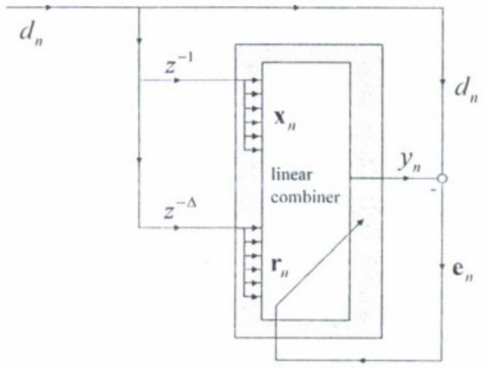

Fig. 1: Adaptive Linear Prediction Scenario.

For both the conventional and the two-channel adaptive filter (AF) the NLMS algorithm is used, implemented as follows.

$$\begin{aligned} e_n &= d_n - \mathbf{w}_n^H \mathbf{u}_n \\ \mathbf{w}_{n+1} &= \mathbf{w}_n + \bar{\mu} \frac{e_n^{*}}{\mathbf{u}_n^H \mathbf{u}_n} \mathbf{u}_n \end{aligned} \tag{7}$$

The difference between the conventional and 2-channel cases lies therefore in the definition that is used for the input vector $\mathbf{u}_n$; $\mathbf{u}_n = \mathbf{r}_n$ as in (5) for the conventional case and $\mathbf{u}_n$ is defined as in (8) for the two-channel case.

$$\mathbf{u}_n = \begin{bmatrix} \mathbf{x}_n \\ \mathbf{r}_n \end{bmatrix} \tag{8}$$

The conventional AF is referred to as AF(0,*M*), while the 2-channel AF is referred to as AF(*L*,*M*).

3. WIENER FILTERS

When s_n and n_n in (1) are wide sense stationary Gaussian processes, as assumed here, the optimal predictor (a Wiener filter) is in fact linear and time-invariant, as in (2) and (3).

For the 2-channel case, the Wiener filter (WF) design follows from the following general Wiener-Hopf equation.

$$E\left\{\begin{bmatrix}\mathbf{x}_n\\ \mathbf{r}_n\end{bmatrix}\begin{bmatrix}\mathbf{x}_n^H & \mathbf{r}_n^H\end{bmatrix}\right\}\mathbf{w}_{xr} = E\left\{\begin{bmatrix}\mathbf{x}_n\\ \mathbf{r}_n\end{bmatrix}d_n^H\right\} \qquad (9)$$

For the conventional WF design we can use (9) after deleting the partition corresponding to the auxiliary channel. Recognizing that the noise-free process s_n is AR(1) and that the noise n_n is white and zero-mean, the component matrices needed in (9) are seen to be auto- or cross-covariances of ARMA processes. The latter can be evaluated using AR [6] and Sylvester matrix based techniques [7] respectively. The performance of the resulting WF is given by

$$MMSE_{WF(L,M)} = E\left\{\left|d_n\right|^2\right\} - E\left\{\begin{bmatrix}\mathbf{x}_n\\ \mathbf{r}_n\end{bmatrix}d_n^*\right\}\mathbf{w}_{xr}^H \qquad (10)$$

Note that the scenarios reflected above are wide-sense stationary, so that all resulting WF solutions correspond to linear time-invariant (LTI) filters.

4. NLMS RESULTS

To illustrate the nonlinear effects that occur when using NLMS in the conventional ALP scenario, we use an AR(1) process with its pole p at $0.95e^{j\pi/6} = 0.823 + j0.475$, additive white noise so that SNR is 80 dB, and aim to predict Δ=10 samples ahead. For the conventional ALP we choose M=2, and for the 2-channel case we choose L=1 and M=2. The conventional WF – denoted WF(0,2) – and the 2-channel WF – denoted WF(1,2) – are designed according to (9), and the performance of each is evaluated using (10). The theoretical minimum mean-square error (MMSE), from (10), is 1 (0 dB) for WF(1,2) and 6.58 (8.18 dB) for WF(0,2). The NLMS algorithm is used with stepsize $\bar{\mu} = 0.01$ for a total of 10,000 iterations, starting with the initial weights set to zero. Convergence was observed to have occurred after 5,000 iterations. Running the above WF designs on this data, yielded typical MSE estimates of 0.978 for WF(1,2) and of 6.79 for WF(0,2). Note that both WF designs are time-invariant in the given scenario. The corresponding NLMS adaptive filters produced MSE estimates (computed from the final 1,000 iterations, i.e. in steady-state) of 0.987 for AF(1,2) and of 6.83 for AF(0,2). Note that in the conventional as well as 2-channel cases the AF and WF results are very close to the theoretical expectation. The real parts of the AF weights during the final 1,000 iterations are given in Figures 2 and 3 (the imaginary parts behave similarly; their plots are not included due to space limitations).

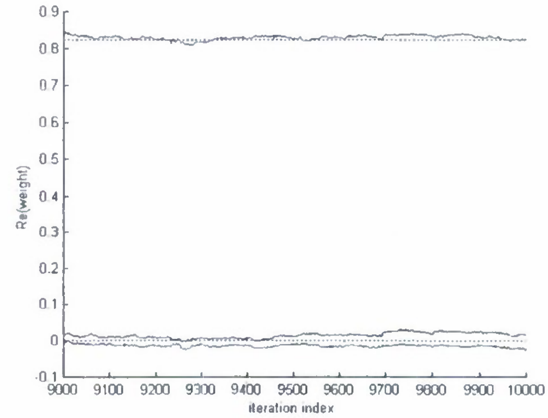

Fig. 2: Re(Weights) AF(1,2) for $\bar{\mu} = 0.01$.

Note that the AF(1,2) weights in Fig. 2 vary slightly, and do so about the TI WF(1,2) weights (dotted constant). The latter are p^*, 0, and 0 (this result from (9) will be made clear in Section 6).

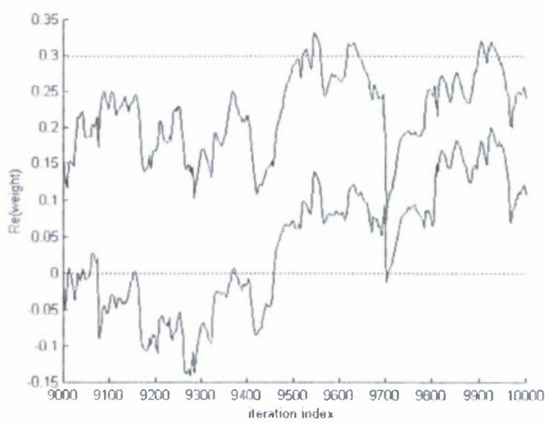

Fig. 3: Re(Weights) AF(0,2) for $\bar{\mu} = 0.01$.

We note that the AF(0,2) weights, i.e. for the conventional AF, vary much more than the AF(1,2) weights. Recall that the MSE performance of 0.987 was fairly close to the theoretically expected performance of 1 for the TI WF, as well as close to the MSE performance of 0.978 of the experimental run of the WF(0,2) as designed. We observe that the AF(0,2) weights vary about the weights of the TI WF(0,2) (dotted constant).

For small stepsize, NLMS puts a premium on finding a constant weight vector, if one exists. The observations above suggest that such a constant solution exists in the 2-channel case. The conventional AF(0,2) weight behavior is not as constant by far, yet it produces performance close to that of the TI WF(0,2). We hypothesize that – for small stepsize – the NLMS AF(0,2) weights remain close to the best long term average, constant solution.

We now repeat the above experiment, with as the basic change that the stepsize is now large, in fact equal to 1. In this situation NLMS puts a premium on tracking time-varying weights, if that is what the structure of the desired process represents. Since NLMS convergence is fastest for this large stepsize, we now run 5,000 iterations, with the final 1,000 iterations designating the steady state region. Note that the theoretically expected MSE performance for WF(1,2) and WF(0,2) is invariant to NLMS stepsize, and therefore remain the same, at 1 and 6.58 respectively. The experimental MSE performance for the designed WF are 1.03 for WF(1,2) and 6.20 for WF(0,2), i.e. close to the expected MSE performance. For the AF operations during steady state we find 2.10 for AF(1,2) and 4.56 for AF(0,2). Note that AF(1,2), for which a time-invariant solution exists, incurs excess MSE. However, AF(0,2) – for which the existence of a TI solution was already doubtful – now produces MSE that is less than the TI WF expectation for MSE performance.

We see that the conventional NLMS AF(0,2) performs better than its TI WF(0,2) counterpart, thereby illustrating nonlinear or non-Wiener behavior. Note that the 2-channel NLMS AF(1,2) performs better than the conventional NLMS AF(0,2) but worse than its TI WF(1,2) counterpart.

The real parts of the AF weights during the final 1,000 iterations are given in Figures 4 and 5.

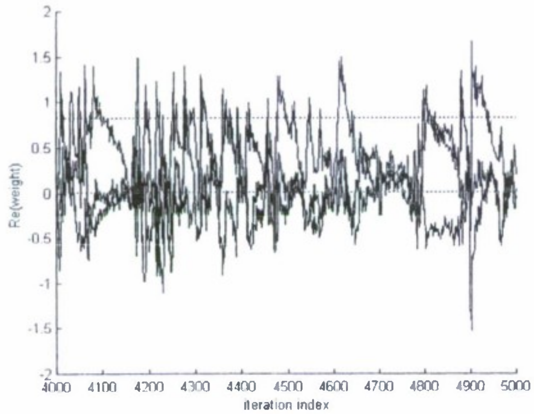

Fig. 4: Re(Weights) AF(1,2) for $\bar{\mu} = 1$.

The AF(1,2) weights now vary substantially, as a result of $\bar{\mu} = 1$, accounting for the large excess MSE.

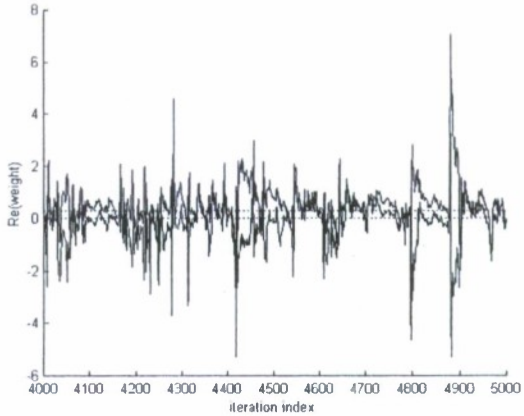

Fig. 5: Re(Weights) AF(0,2) for $\bar{\mu} = 1$.

We also note substantial time-varying behavior of the AF(0,2) weights. In this case this time-varying weight behavior explains the improvement in MSE performance.

Note that when NLMS stepsize is large, there is an immediate a posteriori adjustment of the AF weight vector according to the error signal, as computed in (7). The latter reflects the discrepancy between the desired data and its current model, as reflected in the a priori weight vector. In general, larger stepsizes are good for tracking of time-varying weights, while small stepsizes are good for reducing excess MSE. The results shown in Figs. 2 through 5 suggest that the conventional NLMS AF is in tracking mode (doing better at large stepsize, and exceeding WF performance), while the 2-channel NLMS AF is in estimation mode (doing better at small stepsize, and approaching WF performance). We will elaborate on this finding in Section 6.

The above results were for single realizations of the desired and noise processes. We provide corresponding results for 5 different realizations in Fig. 6.

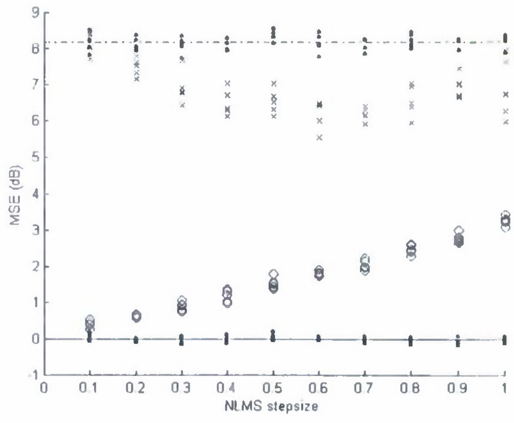

Fig. 6: MSE Performance with $p = 0.95e^{j\pi/6}$.

Fig. 6 shows, at various stepsizes, experimental MSE performance (over 1,000 iterations in steady state) for

WF(0,2) (heavy dots top), AF(0,2) (x's near top), AF(1,2) (o's near bottom), and WF(1,2) (heavy dots bottom). The theoretical MMSE for WF(0,2) (constant line near top) and for WF(1,2) (constant line near bottom) are also indicated. Note that the maximum MSE performance improvement achieved by AF(0,2) – over WF(0,2) – is approximately 2.5 dB (seen at stepsize 0.6). The optimal performance – approached by AF(1,2) at small stepsize – is 5.5 dB better still.

The AF(1,2) performance is at the WF(1,2) performance for small stepsize, and linearly worsens as stepsize increases, as typical for a time-invariant structure underlying the desired data. The AF(0,2) performance is also at the WF(0,2) performance for small stepsize, but shows performance improvements for increased stepsizes (perhaps saturating at some level, or worsening for very high stepsize), as indicative of a time-varying structure underlying the desired data.

The next section elaborates on the origin of the time-varying behavior when using AF(0,2), i.e. the conventional AF.

6. TV WF SOLUTION

For the above scenarios, at SNR=80 dB, the desired process is almost purely AR(1). Consequently, the process d_n pretty much satisfies the following structure.

$$d_n = p\, d_{n-1} + \varepsilon_n \tag{11}$$

The optimal Δ-step estimator for an AR(1) process is given by

$$\hat{d}_n = p^{\Delta}\, d_{n-\Delta} \tag{12}$$

so that we recognize the first RHS term in (11) as the optimal one-step estimator. MSE performance deteriorates as Δ is increased, so that the one-step estimator delivers the best possible performance. Writing the optimal estimator in the form of the 2-channel AF(1,2) model yields the following (where the variance of ε_n is minimal).

$$\begin{aligned} d_n &= \hat{d}_n + \varepsilon_n \\ &= p\, d_{n-1} + \varepsilon_n \\ &= \begin{bmatrix} p^* & 0 & 0 \end{bmatrix}^H \begin{bmatrix} d_{n-1} \\ d_{n-\Delta} \\ d_{n-\Delta-1} \end{bmatrix} + \varepsilon_n \\ &= \mathbf{w}_{xr}^H \mathbf{u}_n + \varepsilon_n \end{aligned} \tag{13}$$

The latter shows explicitly which TI weight vector solution AF(1,2) aims for.

In order to rewrite the desired data structure in terms of the conventional AF model, we introduce linking sequences, connecting the auxiliary channel element to the reference channel elements.

$$\begin{aligned} \rho_{n-\Delta}^{(\Delta-1)} &= \frac{d_{n-1}}{d_{n-\Delta}} \\ \rho_{n-\Delta-1}^{(\Delta)} &= \frac{d_{n-1}}{d_{n-\Delta-1}} \end{aligned} \tag{14}$$

Using the linking sequences in (14) we substitute for the auxiliary channel element in (13). Making the substitution in terms of $d_{n-\Delta}$ and, alternatively, in terms of $d_{n-\Delta-1}$, and recognizing that each is equally valid, taking an affine linear combination of the result yields the following equivalent structure for the desired process.

$$\begin{aligned} d_n &= \begin{bmatrix} p^* & 0 & 0 \end{bmatrix}^H \begin{bmatrix} d_{n-1} \\ d_{n-\Delta} \\ d_{n-\Delta-1} \end{bmatrix} + \varepsilon_n \\ &= \alpha\, p\, \rho_{n-\Delta}^{(\Delta-1)} d_{n-\Delta} + (1-\alpha)\, p \rho_{n-\Delta-1}^{(\Delta)} d_{n-\Delta-1} + \varepsilon_n \\ &= p^* \begin{bmatrix} \alpha \rho_{n-\Delta}^{(\Delta-1)*} & (1-\alpha)\rho_{n-\Delta-1}^{(\Delta)*} \end{bmatrix}^H \begin{bmatrix} d_{n-\Delta} \\ d_{n-\Delta-1} \end{bmatrix} + \varepsilon_n \\ &= \mathbf{w}_{0r,n}^H \mathbf{r}_n + \varepsilon_n \end{aligned} \tag{15}$$

The latter equation shows that there is a target weight vector for NLMS AF(0,2), the conventional AF, corresponding to the same MMSE as that of the one-step predictor. However, the latter MMSE can only be realized if AF(0,2) can faithfully track the following time-varying weight vector implied by (15).

$$\mathbf{w}_{0r,n} = p \begin{bmatrix} \alpha \rho_{n-\Delta}^{(\Delta-1)*} \\ (1-\alpha)\rho_{n-\Delta-1}^{(\Delta)*} \end{bmatrix} \tag{16}$$

The AR(1) process in our examples dictates the following linking sequence behavior.

$$\begin{aligned} \rho_{n-\Delta}^{(\Delta-1)} &= \frac{d_{n-1}}{d_{n-\Delta}} = p^{\Delta-1} + \upsilon_{n-\Delta} \\ \rho_{n-\Delta-1}^{(\Delta)} &= \frac{d_{n-1}}{d_{n-\Delta-1}} = p^{\Delta} + \eta_{n-\Delta-1} \end{aligned} \tag{17}$$

Note that upon substituting the constant component from (17) into (16), the constant portion of that weight vector is recognized to consist of an affine combination of the optimal Δ-step and (Δ+1)-step predictors given in (12). The stochastic components of the linking sequences in (17) – which in our scenario arises from the driving noise of the AR(1) process – cause the AF(0,2) target weight vector to be time-varying about the above constant portion. As a result, AF(0,2) incurs not only an estimation error but also a tracking error, and its performance is not as good as that of WF(1,2). However, partial success in tracking explains performance improvement over that of WF(0,2).

The latter observation suggests that the tracking error would be smaller if the stochastic components in (17) were smaller. The driving noise is smaller, relative to the desired process (and away from its zero-crossings), when the AR(1) process is more narrowband. Rerunning the experiment that produced Fig. 6, but now for $p = 0.99e^{j\pi/6}$, yields the results in Fig. 7.

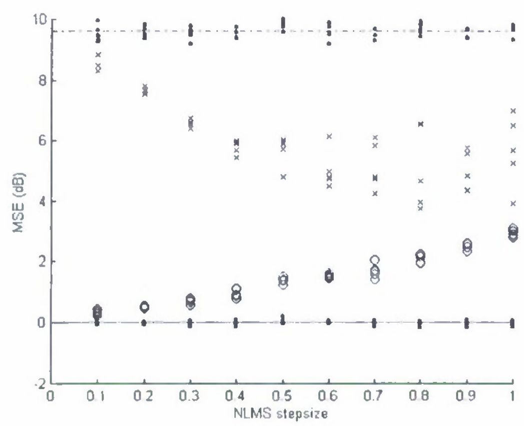

Fig. 7: MSE Performance with $p = 0.99e^{j\pi/6}$.

While the overall behavior in Fig. 7 is similar to that in Fig. 6, we note that the desired signal variance has increased (from 8.18 dB to 9.61 dB above the driving noise variance of 1) while the best AF(0,2) MSE performance is now approximately 6 dB better than for WF(0,2) (seen at stepsize 0.8). Note that the latter performance improvement is not only larger than for the wider bandwidth process used for Fig. 6, it is now also to within 4 dB of the possible optimal performance.

7. CONCLUSION

We have shown that the nonlinear effects of adaptive filtering in the linear prediction scenario are associated with time-varying NLMS weights. Based on the 2-channel ALP scenario an expression for the time-varying target weight vector was given, which the conventional NLMS adaptive filter aims to track. The time-varying nature corresponds to the structure underlying the desired data. The conventional NLMS adaptive filter can outperform the time-invariant Wiener filter of the same filter order and is bounded in its performance by that of the optimal 2-channel Wiener filter.

8. ACKNOWLEDGEMENT

The present work was funded by the Independent Laboratory Research Program at SPAWAR Systems Center, San Diego, and – in part – by the National Research Council, when the first author was a Senior Research Associate at SPAWAR Systems Center, San Diego, during his Fall 2001 sabbatical semester from Virginia Tech.

REFERENCES

[1] M. Reuter, K. Quirk, J. Zeidler, and L. Milstein, Nonlinear effects in LMS adaptive filters, *Proc. Symp. 2000 on Adaptive Systems for Signal Processing, Communications and Control*, 141-146, Lake Louise, Alberta, October 2000.

[2] K. J. Quirk, L. B. Milstein, and J. R. Zeidler, "A performance bound of the LMS estimator," IEEE Trans. Information Theory, 46, 1150-1158, May 2000.

[3] M. Reuter and J. R. Zeidler, Nonlinear effects in LMS adaptive equalizers, *IEEE Trans. Signal Processing*, 47, 1570-1579, June 1999.

[4] J. Han, J. R. Zeidler, and W. H. Ku, Nonlinear Effect of the LMS Adaptive Predictor for Chirped Input Signals, *EUROSIP Journal on Applied Signal Processing*, Special Issue on Nonlinear Signal and Image Processing, Part II, pp. 21-29, January 2002.

[5] A. A. (Louis) Beex and James R. Zeidler, "Nonlinear Effects in Adaptive Filters," in Advances in LMS Adaptive Filters, eds. S. Haykin and B. Widrow, John Wiley & Sons, 2002.

[6] J-P. Dugré, A. A. (Louis) Beex, and L. L. Scharf, Generating Covariance Sequences and the Calculation of Quantization and Rounding Error Variances in Digital Filters, *IEEE Trans. Acoustics, Speech, and Signal Processing*, 28, 102-104, February 1980.

[7] A. A. (Louis) Beex, Efficient Generation of ARMA Cross-Covariance Sequences, *Proc. Int'l Conf. On Acoustics, Speech, and Signal Processing*, 327-330, Tampa FL, 26-29 March 1985.